Asymptotic analysis of discounted zero-sum games: some recent advances

Sylvain Sorin

Asymptotic analysis of discounted zero-sum games: some recent advances

Sylvain Sorin

IMJ
Université P. et M. Curie - Paris 6

Conference “New Trends in Optimal Control”
Tours, June 23-27, 2014

This is a joint work with Guillaume Vigeral

Table of contents

Table of contents

A finite two person zero-sum stochastic game is defined by a state space Ω, actions spaces I and J, a transition probability Q from $\Omega \times I \times J$ to $\Delta(\Omega)$ and a real payoff function g on $\Omega \times I \times J$.
Time is discrete: at stage n given the past history including ω_n the players choose (at random) i_n and j_n, the stage payoff is $g_n = g(\omega_n, i_n, j_n)$ and the law of the new state ω_{n+1} is $q(\omega_n, i_n, j_n)$.
For a λ-discounted evaluation of the payoff (Shapley (1953)), one obtains existence of the value v_λ, unique solution of :

$$v_\lambda(\omega) = \mathtt{val}_{X \times Y}[\lambda g(\omega, i, j) + (1-\lambda) \sum_{\omega'} q(\omega, i, j)(\omega') v_\lambda(\omega')]$$

where $X = \Delta(I), Y = \Delta(J)$.
This **recursive formula** extends to :
1) general repeated games (incomplete information, signals ...)
2) general evaluation, defined by a probability $\{\theta_n\}_{n \geq 1}$ and then $g = \sum_n \theta_n g_n$.

A finite two person zero-sum stochastic game is defined by a state space Ω, actions spaces I and J, a transition probability Q from $\Omega \times I \times J$ to $\Delta(\Omega)$ and a real payoff function g on $\Omega \times I \times J$.
Time is discrete: at stage n given the past history including ω_n the players choose (at random) i_n and j_n, the stage payoff is $g_n = g(\omega_n, i_n, j_n)$ and the law of the new state ω_{n+1} is $q(\omega_n, i_n, j_n)$.
For a λ-discounted evaluation of the payoff (Shapley (1953)), one obtains existence of the value v_λ, unique solution of :

$$v_\lambda(\omega) = \mathtt{val}_{X \times Y}[\lambda g(\omega, i, j) + (1-\lambda) \sum_{\omega'} q(\omega, i, j)(\omega') v_\lambda(\omega')]$$

where $X = \Delta(I), Y = \Delta(J)$.
This **recursive formula** extends to :
1) general repeated games (incomplete information, signals ...)
2) general evaluation, defined by a probability $\{\theta_n\}_{n \geq 1}$ and then $g = \sum_n \theta_n g_n$.

The **asymptotic analysis** is the study of the sequence of values along a family of evaluations with mean going to ∞.
Example of proof: algebraic approach for finite discounted stochastic games, Bewley-Kohlberg (1976).
Open problems:
i) stochastic games: finite state space and compact actions space,
ii) general finite repeated game.
Recents counter examples cover:
Case i) Guillaume Vigeral (2013)
Case ii) Bruno Ziliotto (2013): finite action spaces and public signals on the state space.
We show that by coupling two stopping time problems one can construct two-person zero-sum stochastic games with finite state space having oscillating discounted values.
This unifies and generalizes the above recent results.

Table of contents

Configuration

A *configuration* P is defined by a general zero-sum repeated game on a state space $\Omega = \overline{\Omega} \bigcup \{\bar{\omega}\}$ starting from some state $\omega \in \Omega$, known to the players.

Each couple of strategies (σ, τ) of the players specifies, with the parameters of the game, the law of the stopping time S of exit of $\overline{\Omega}$:

$$S = \min\{n \in \boldsymbol{N}; \omega_n \notin \overline{\Omega}\}.$$

For each evaluation $\theta = \{\theta_n\}$ on $\boldsymbol{N}^* = 1, 2, ...$, let $g_\theta(\sigma, \tau)$ be the expected (normalized) duration spent in $\overline{\Omega}$.

$$g_\theta(\sigma, \tau) = E_{\sigma,\tau}[\sum_{n=1}^{S-1} \theta_n].$$

Assume that Player 1 (resp 2) wants to minimize (resp maximize) this quantity, and define the inertia rate:

$$Q_\theta = \sup_\tau \inf_\sigma g_\theta(\sigma, \tau) = \inf_\sigma \sup_\tau g_\theta(\sigma, \tau)$$

Then for each $\alpha < \beta$, in the game with evaluation θ and payoff α in $\overline{\Omega}$ and β in the absorbing state $\bar{\omega}$, Player 1 (the maximizer) minimizes $g_\theta(\sigma, \tau)$ since the payoff is given by

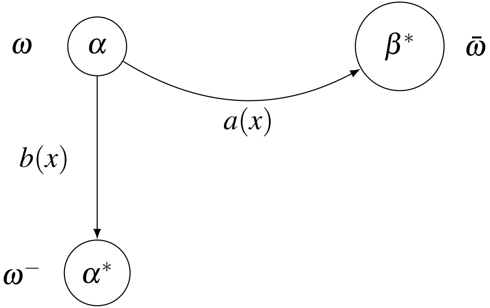
$$\gamma_\theta(\sigma, \tau) = \alpha g_\theta(\sigma, \tau) + \beta(1 - g_\theta(\sigma, \tau)).$$

In particular if the game has a value v_θ then

$$v_\theta = \alpha Q_\theta + \beta(1 - Q_\theta)$$

A first simple example is a dynamic programming problem P. with three states $\Omega = \{\omega, \bar{\omega}, \omega^-\}$ where both $\bar{\omega}$ and ω^- are absorbing.

The action set is $X = [0,1]$ and the impact of an action x is on the transitions, given by $a(x)$ from ω to $\bar{\omega}$ and $b(x)$ from ω to ω^-, where a and b are two continuous function from $[0,1]$ to $[0,1]$.



Here for any discount factor λ,

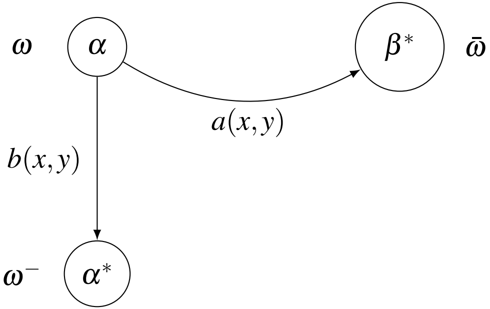
$$Q_\lambda = \inf_x \frac{\lambda + (1-\lambda) b(x)}{\lambda + (1-\lambda) b(x) + (1-\lambda) a(x)}$$

Another example consist of half part of a game defined by Bewley and Kohlberg (1978).
This is a 3 states example with 2 absorbing states: $\bar{\omega}$ with payoff β, ω^- with payoff α, and $\beta > \alpha$. In state ω the payoff is α and the transitions are given by:

	Stay	Quit
Stay	ω	$\bar{\omega}$
Quit	$\bar{\omega}$	ω^-

Let x (resp. y) be the probability on Stay for player 1 (resp. 2).

The mixed extension is the *configuration*:



with $a(x,y) = x(1-y) + y(1-x),\ b(x,y) = xy$.
Here for any discount factor λ,

$$
\begin{aligned}
Q_\lambda &= \inf_x \sup_y \frac{\lambda + (1-\lambda)b(x,y)}{\lambda + (1-\lambda)b(x,y) + (1-\lambda)a(x,y)} \\
&= \inf_x \sup_y \frac{\lambda + (1-\lambda)b(x,y)}{\lambda + (1-\lambda)b(x,y) + (1-\lambda)a(x,y)}
\end{aligned}
$$

Table of contents

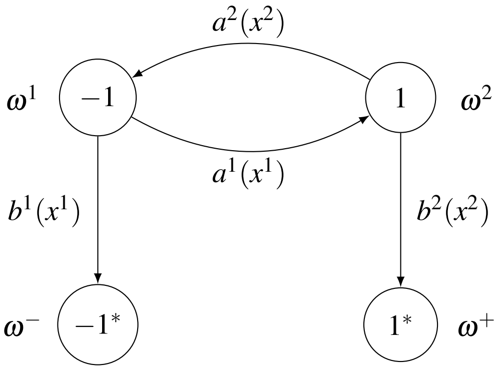
Reversibility

Consider now a two person zero-sum stochastic game G generated by two dual configurations P^1 and P^2 of the previous type, with $\alpha^1 = -1$, $\alpha^2 = 1$, which are coupled:
the exit domain from P^1 ($\Omega^1 \setminus \overline{\Omega}^1$) is the starting state ω^2 in P^2 and reciprocally.

For example

Reversibility

Consider now a two person zero-sum stochastic game G generated by two dual configurations P^1 and P^2 of the previous type, with $\alpha^1 = -1$, $\alpha^2 = 1$, which are coupled:
the exit domain from P^1 ($\Omega^1 \setminus \overline{\Omega}^1$) is the starting state ω^2 in P^2 and reciprocally.
For example



In addition we assume that the exit event is known by the players and that the payoff are discounted.
For each $\lambda \in]0,1]$ this defines a game G_λ with value v_λ satisfying:
$v_\lambda(\omega^1) = v_\lambda^1 \in [-1,1[, v_\lambda(\omega^2) = v_\lambda^2 \in]-1,1]$.
In particular, starting from state ω^1 the model is equivalent to the one with an absorbing exit state ω^2 with payoff $v_\lambda(\omega^2)$ (by stationarity), which thus corresponds to the payoff β in the configuration P of the previous section.

Proposition

$$v_\lambda^1 = Q_\lambda^1 \times (-1) + (1 - Q_\lambda^1) \times v_\lambda^2$$
$$v_\lambda^2 = Q_\lambda^2 \times (+1) + (1 - Q_\lambda^2) \times v_\lambda^1.$$

Corollary

$$v_\lambda^1 = \frac{Q_\lambda^2 - Q_\lambda^1 - Q_\lambda^1 Q_\lambda^2}{Q_\lambda^1 + Q_\lambda^2 - Q_\lambda^1 Q_\lambda^2}$$
$$v_\lambda^2 = \frac{Q_\lambda^2 - Q_\lambda^1 + Q_\lambda^1 Q_\lambda^2}{Q_\lambda^1 + Q_\lambda^2 - Q_\lambda^1 Q_\lambda^2}$$

Comments

1) As λ goes to 0, Q_λ converges to 0 (and v_λ converges to β) in the configuration of Section 1, as soon as $\limsup \frac{a(x)}{b(x)} = +\infty$, as x goes to 0.

2) In the current framework, assuming that both Q^i_λ go to 0, the asymptotic behavior of v^1_λ depends upon the evolution of the ratio $\frac{Q^1_\lambda}{Q^2_\lambda}$. In fact one has:

$$v^1_\lambda \sim v^2_\lambda \sim \frac{1 - \frac{Q^1_\lambda}{Q^2_\lambda}}{1 + \frac{Q^1_\lambda}{Q^2_\lambda}}.$$

Theorem

Assume that both Q^i_λ go to 0 and that $\frac{Q^1_\lambda}{Q^2_\lambda}$ has more than one accumulation point, then v^i_λ does not converge.

Comments

1) As λ goes to 0, Q_λ converges to 0 (and v_λ converges to β) in the configuration of Section 1, as soon as $\limsup \frac{a(x)}{b(x)} = +\infty$, as x goes to 0.
2) In the current framework, assuming that both Q^i_λ go to 0, the asymptotic behavior of v^1_λ depends upon the evolution of the ratio $\frac{Q^1_\lambda}{Q^2_\lambda}$. In fact one has:

$$v^1_\lambda \sim v^2_\lambda \sim \frac{1 - \frac{Q^1_\lambda}{Q^2_\lambda}}{1 + \frac{Q^1_\lambda}{Q^2_\lambda}}.$$

Theorem

Assume that both Q^i_λ go to 0 and that $\frac{Q^1_\lambda}{Q^2_\lambda}$ has more than one accumulation point, then v^i_λ does not converge.

Comments

1) As λ goes to 0, Q_λ converges to 0 (and v_λ converges to β) in the configuration of Section 1, as soon as $\limsup \frac{a(x)}{b(x)} = +\infty$, as x goes to 0.
2) In the current framework, assuming that both Q^i_λ go to 0, the asymptotic behavior of v^1_λ depends upon the evolution of the ratio $\frac{Q^1_\lambda}{Q^2_\lambda}$. In fact one has:

$$v^1_\lambda \sim v^2_\lambda \sim \frac{1-\frac{Q^1_\lambda}{Q^2_\lambda}}{1+\frac{Q^1_\lambda}{Q^2_\lambda}}.$$

Theorem

Assume that both Q^i_λ go to 0 and that $\frac{Q^1_\lambda}{Q^2_\lambda}$ has more than one accumulation point, then v^i_λ does not converge.

Comments

More precisely it is enough that $Q^i_\lambda \sim \lambda^r c^i(\lambda)$ for some $r > 0$, whith $0 < A \leq c^i \leq B$ and that one of the c^i does not converge to obtain the result.
The next section will describe several configurations generating such probabilities Q^i_λ, with c^i converging or not.
We will use the terminology *regular or oscillating configurations*.

Table of contents

Example A0 : Regular with 0 player

Consider a random walk on $\mathbf{N} \cup \{-1\}$ and exit state -1. In any other state $m \in \mathbf{N}$ the transition is $\frac{1}{2}\delta_{m-1} + \frac{1}{2}\delta_{m+1}$. The starting state is 0. Denote by s_n the probability that exit happens at stage n ; it is well known that the generating function of S is given by $F(z) = \frac{1-\sqrt{1-z^2}}{z}$. Hence,

$$
\begin{aligned}
Q_\lambda &= \sum_{n=1}^{+\infty} s_n \sum_{t=1}^{n} \lambda (1-\lambda)^{i-1} \\
&= \sum_{n=1}^{+\infty} s_n (1-(1-\lambda)^n) \\
&= F(1) - F(1-\lambda) \\
&= \frac{\sqrt{2\lambda - \lambda^2} - \lambda}{1-\lambda} \\
&\sim \sqrt{2\lambda}.
\end{aligned}
$$

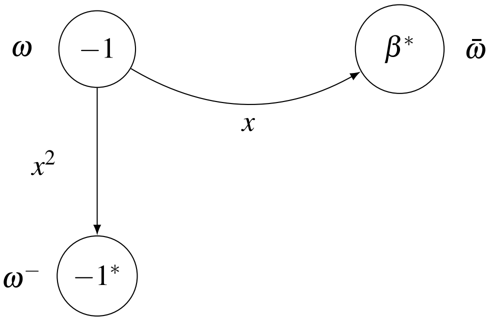
Example A1 : Regular with one player

Take $a(x) = x$ and $b(x) = x^2$.
Then x_λ minimizes $g_\lambda(x)$ with

$$g_\lambda(x) = \frac{(\lambda + (1-\lambda)b(x))}{\lambda + (1-\lambda)(a(x)+b(x))} \tag{1}$$

so that: $x_\lambda^2 = \frac{\lambda}{1-\lambda}$ and

$$Q_\lambda \sim 2\sqrt{\lambda}. \tag{2}$$

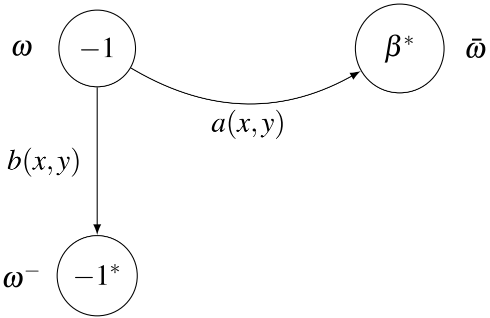


Example A2: regular configuration with 2 players

BK- configuration:

	Stay	Quit
Stay	ω	$\bar{\omega}$
Quit	$\bar{\omega}$	ω^-

x (resp. y) is the probability on Stay and
$a(x,y) = x(1-y) + y(1-x)$, $b(x,y) = xy$.



Proposition

$$Q_\lambda \sim \sqrt{\lambda}.$$

One has

$$g_\lambda(x,y) = \frac{\lambda + (1-\lambda)b(x,y)}{\lambda + (1-\lambda)b(x,y) + (1-\lambda)a(x,y)}$$

.

Hence $x_\lambda = y_\lambda \sim \sqrt{\lambda}$.
Thus this configuration induces exit transitions $a(x_\lambda, y_\lambda)$ of the order of $2\sqrt{\lambda}$ and $b(x_\lambda, y_\lambda)$ of the order of λ, so that $Q_\lambda \sim \sqrt{\lambda}$.
This defines a regular configuration similar to example A1.

Example B1: Example A1 perturbed

To get oscillations one can choose:
$a(x) = xf(x)$ with $f(x)$ bounded away from 0, oscillating and such that $f'(x) = o(1/x)$.
For example, $f(x) = 2 + \sin(ln(-lnx))$ and $b(x) = x^2$.

Proposition

For this choice of transition functions one has:

$$Q_\lambda \sim \frac{2\sqrt{\lambda}}{f(\sqrt{\lambda})}.$$

Example B2: example A2 perturbed

Let $s \in C^1([0, \frac{1}{16}], \mathbb{R})$ such that s and $x \to xs'(x)$ are both bounded by $\frac{1}{16}$. Consider a configuration as before but for perturbed functions a and b:

$$\begin{aligned} a(x,y) &= \frac{(\sqrt{x}+\sqrt{y})(1-\sqrt{x}+s(x))(1-\sqrt{y}+s(y))}{2(1-x)(1-y)(1-f_2(x,y))} \\ b(x,y) &= \frac{\sqrt{xy}\left[(1-\sqrt{x})(1-\sqrt{y})+f_1(x,y)-\sqrt{xy}f_2(x,y)\right]}{(1-x)(1-y)(1-f_2(x,y))}. \end{aligned}$$

where

$$f_1(x,y) = \begin{cases} \frac{\sqrt{x}s(x)-\sqrt{y}s(y)}{\sqrt{x}-\sqrt{y}} & \text{if } x \neq y \\ 2xs'(x)+s(x) & \text{if } x = y \end{cases}$$

and

$$f_2(x,y) = \begin{cases} \frac{\sqrt{y}s(x)-\sqrt{x}s(y)}{\sqrt{x}-\sqrt{y}} & \text{if } x \neq y \\ 2xs'(x)-s(x) & \text{if } x = y \end{cases}$$

Example B2: example A2 perturbed

Let $s \in C^1([0, \frac{1}{16}], \mathbb{R})$ such that s and $x \to xs'(x)$ are both bounded by $\frac{1}{16}$. Consider a configuration as before but for perturbed functions a and b:

$$\begin{aligned} a(x,y) &= \frac{(\sqrt{x}+\sqrt{y})(1-\sqrt{x}+s(x))(1-\sqrt{y}+s(y))}{2(1-x)(1-y)(1-f_2(x,y))} \\ b(x,y) &= \frac{\sqrt{xy}\left[(1-\sqrt{x})(1-\sqrt{y})+f_1(x,y)-\sqrt{xy}f_2(x,y)\right]}{(1-x)(1-y)(1-f_2(x,y))}. \end{aligned}$$

where

$$f_1(x,y) = \begin{cases} \frac{\sqrt{x}s(x)-\sqrt{y}s(y)}{\sqrt{x}-\sqrt{y}} & \text{if } x \neq y \\ 2xs'(x)+s(x) & \text{if } x = y \end{cases}$$

and

$$f_2(x,y) = \begin{cases} \frac{\sqrt{y}s(x)-\sqrt{x}s(y)}{\sqrt{x}-\sqrt{y}} & \text{if } x \neq y \\ 2xs'(x)-s(x) & \text{if } x = y \end{cases}$$

Then by Vigeral (2013) these functions a and b are continuous, and one can compute that the configuration is oscillating of order $\frac{1}{2}$.

The next 4 configurations are an outcome of the work of Ziliotto (2013) and use exactly the same tools.

Example B3: Countable action space

We consider again example A1 but now the action space X is $\boldsymbol{N}^*$ and the transition are given by $(a_n, b_n) = (\frac{1}{2^n}, \frac{1}{4^n})$.

Proposition

For this configuration $Q_\lambda/\sqrt{\lambda}$ oscillates on a sequence $\{\lambda_m\}$ of discount factors like $\lambda_m = \frac{1}{2^m}$.

Note first that the choice n inducing $x = \frac{1}{2^n}$ is asymptotically optimal (recall $a(x) = x, b(x) = x^2$), at $\lambda_{2n} = \frac{1}{4^n}$ and thus $Q_\lambda \sim 2\sqrt{\lambda}$.
For $\lambda = \lambda_{2n+1}$ one obtains:

$$g_\lambda(\frac{1}{2^n}) \sim (\frac{1}{2} \times \frac{1}{4^n} + \frac{1}{4^n})2^n \sim \frac{3\sqrt{2}}{2}\sqrt{\lambda}$$

$$g_\lambda(\frac{1}{2^{n+1}}) \quad \sim \quad (\frac{1}{2} \times \frac{1}{4^n} + \frac{1}{4^{n+1}})2^{n+1} \sim \frac{3\sqrt{2}}{2}\sqrt{\lambda}$$

Finally one checks that

$$g_\lambda(\frac{1}{2^{n+m}}) \geq \frac{3\sqrt{2}}{2}\sqrt{\lambda}$$

for $m = -n, ..., -1$ and $m \geq 2$.
Thus for this specific λ, $g_\lambda(x)$ is bounded below by a quantity of the order $\frac{3\sqrt{2}}{2}\sqrt{\lambda}$.
It follows that $Q_\lambda/\sqrt{\lambda}$ oscillates between 2 and $\frac{3\sqrt{2}}{2}$ on a sequence $\{\lambda_m\}$ of discount factors like $\lambda_m = \frac{1}{2^m}$.

Exemple B4 :Countable state space

We consider here a configuration which is the dual of the previous one with finite action space and countably many states.
The state space countains a countable subspace of probabilities $y = (y^A, y^B)$ on two positions A and B with $y_n = (\frac{1}{2^n}, 1 - \frac{1}{2^n}), n = 0, 1, ...$, and two absorbing states A^* and B^*.
Hence $\Omega = \{y_n, n \in \boldsymbol{N}; A^*, B^*\}$, $\omega = y_0$ and $\Omega \setminus \overline{\Omega} = \{B^*\}$.
The player has two actions: $Stay$ or $Quit$.
Consider state y_n. Under $Quit$ an absorbing state is reached: A^* with probability y_n^A and B^* with probability y_n^B.
Under $Stay$ the state evolves from y_n to y_{n+1} with probability $1/2$ and to $y_0 = (1, 0)$ with probability $1/2$.
The player is informed upon the state.
A (stationary) strategy of the player can be identified with a stopping time corresponding to the first state y_n when he chooses $Quit$.

Let T_n be the random time corresponding to the first occurrence of y_n (under *Stay*) and θ_n the associated strategy: *Quit* (for the first time) at y_n.

Proposition

$$g_\lambda(\theta_n) = 1 - \frac{(1-\lambda^2)\left(1-\frac{1}{2^n}\right)}{1+2^{n+1}\lambda(1-\lambda)^{-n}-\lambda}$$

Proof.

Lemma 2.5 in Ziliotto (2013) gives

$$\mathsf{E}[(1-\lambda)^{T_n}] = \frac{1-\lambda^2}{1+2^{n+1}\lambda(1-\lambda)^{-n}-\lambda}$$

$\square$

Hence $Q_\lambda = 1-(1-\lambda^2)\max_{n\in\mathbf{N}} \frac{1-\frac{1}{2^n}}{1+2^{n+1}\lambda(1-\lambda)^{-n}-\lambda}$ and one shows, using Ziliotto's Lemma 2.8, that the configuration is irregular : $\frac{Q_\lambda}{\sqrt{\lambda}}$ oscillates between two positive values.

Example B5: A MDP with signals

The next configuration corresponds to a Markov decision process with 2 states: A and B. The player has 2 actions: *Stay* or *Quit*. The transition are as follows:

A	$\frac{1}{2};\ell$	$\frac{1}{2};r$
Stay	A	$(\frac{1}{2}A+\frac{1}{2}B)$
Quit	A^*	A^*

B	$\frac{1}{2};\ell$	$\frac{1}{2};r$
Stay	A	B
Quit	B^*	B^*

The transition is random: with probability $1/2$ of type ℓ and probability $1/2$ of type r. The player is not informed upon the state reached but only on the signal ℓ or r. The natural “auxiliary state" space is then the beliefs of the player on (A,B) and the model is equivalent to the previous one. In fact under Stay, ℓ occurs with probability $1/2$ and the new parameter is $y_0=(1,0)$; and after r, the belief evolves from y_n to y_{n+1}. Again this configuration generates an oscillating Q_λ of the order of $\sqrt{\lambda}$.

Example B6: A symmetric game with no signal on the state

A next transformation is to introduce two players and to generate the random variable $\frac{1}{2}(\ell) + \frac{1}{2}(r)$ in the above model through the moves of the players (joint controlled lottery).
This leads to the original framework of the game defined by Ziliotto (2013): action and state spaces are finite and the only information of the players is the sequence of moves and the initial state $\omega = A$.
Here $\Omega = \{A, B, A^*, B^*\}$ and $\Omega \setminus \overline{\Omega} = \{B^*\}$.
Player 1 has three moves: Stay1, Stay2 and Quit, and player 2 has 2 moves: $Left$ and $Right$. The transition are as follows:

A	$Left$	$Right$
Stay1	A	$(\frac{1}{2}A + \frac{1}{2}B)$
Stay2	$(\frac{1}{2}A + \frac{1}{2}B)$	A
Quit	A^*	A^*

B	$Left$	$Right$
Stay1	A	B
Stay2	B	A
Quit	B^*	B^*

By playing $(1/2, 1/2, 0)$ (resp. $(1/2, 1/2)$) player 1 (resp. player 2) can mimick the previous distribution on (ℓ, r) where ℓ corresponds to the event “the moves are on the main diagonal". It follows that this behavior is consistent with optimal strategies hence the induced distribution on plays is like in the previous example B5.

Table of contents

By combining some of these regular/oscillating configurations, one recover several recent examples where v_λ does not converge:

- An example with finitely many states, compact action sets and smooth transitions (Vigeral 2013)
- An example with finitely many states, countable action sets and perfect information (Ziliotto 2013)
- Several examples of finite games with signals or in the dark (Ziliotto 2013)

One also get new examples :

- An example with finitely many states, compact action sets, smooth transitions and perfect information.
- An example with finitely many states, compact action set for Player 1 and finite for Player 2, and smooth transitions.
- Smooth MDP with a countable number of states and 2 actions.

Table of contents

Alternative evaluations

The stationarity of the evaluation (discounted game) was crucial in the computations. However it is possible to construct similar examples in which $\lim v_n$ does not exist.
The idea grounds on a lemma of Neyman (2003) giving sufficient conditions for the two sequences v_n and v_{λ_n}, with $\lambda_n = \frac{1}{n}$, to have the same asymptotic behavior as n tends to infinity.

Irreversible case

The above analysis shows that oscillations in the exit probability and reversibility allows for non convergence of the discounted values.

In fact without reversibility one has convergence, for example in absorbing games or games with incomplete information.

Also notice that any configuration is by itself converging, even when oscillating (Q_λ still tends to 0). It is the association of a regular and an oscillating configuration that implies divergence of v_λ.

A way to forbid oscillations seems to consider semi-algebraic games or more generally games definable in some o-minimal structures (Bolte, Gaubert, Vigeral, 2013).

Table of contents

R. J. Aumann and M. Maschler (1995),
Repeated Games with Incomplete Information,
M.I.T. Press.

T. Bewley and E. Kohlberg (1976),
The asymptotic theory of stochastic games.
Mathematics of Operations Research, **1**, 197-208.

T. Bewley and E. Kohlberg (1976),
The asymptotic solution of a recursion equation occurring in stochastic games.
Mathematics of Operations Research, **1**, 321-336.

T. Bewley and E. Kohlberg (1978),
On stochastic games with stationary optimal strategies.
Mathematics of Operations Research, **3**, 104-125.

J. Bolte, S. Gaubert and G. Vigeral (2013),
Definable zero-sum stochastic games.
http://arxiv.org/abs/1301.1967 .

P. Cardaliaguet, R. Laraki and S. Sorin (2012),

A continuous time approach for the asymptotic value in two-person zero-sum repeated games.
SIAM Journal on Control and Optimization, **50**, 1573-1596.

G. Grimmett and D. Stirzaker (2001),
Probability and Random Processes.
Oxford University Press

E. Kohlberg (1974),
Repeated games with absorbing states,
Annals of Statistics, **2**, 724-738.

J.-F. Mertens, A. Neyman and D. Rosenberg (2009),
Absorbing games with compact action spaces.
Mathematics of Operation Research, **34**, 257-262.

J.-F. Mertens, S. Sorin and S. Zamir (1994),
Repeated Games.
CORE DP 9420-22.

A. Neyman (2003),
Stochastic games and nonexpansive maps,

Chapter 26 in A. Neyman and S. Sorin (eds), *Stochastic Games and Applications*, Kluwer Academic Publishers.

D. Rosenberg and S. Sorin (2001),
An operator approach to zero-sum repeated games,
Israel Journal of Mathematics, **121**, 221-246.

L. S. Shapley (1953),
Stochastic games,
Proceedings of the National Academy of Sciences of the U.S.A., **39**, 1095-1100.

S. Sorin (2003),
The operator approach to zero-sum stochastic games,
Chapter 27 in A. Neyman and S. Sorin (eds), *Stochastic Games and Applications*, Kluwer Academic Publishers.

S. Sorin and G. Vigeral (2013),
Existence of the limit value of two person zero-sum discounted repeated games via comparison theorems,
Journal of Opimization Theory and Applications **157** , 564-576.

S. Sorin and G. Vigeral (2013),
Reversibility and oscillations in zero-sum discounted stochastic games,
hal- 00869656.

G. Vigeral (2013),
A zero-sum stochastic game with compact action sets and no asymptotic value,
Dynamic Games and Applications, **3**, 172-186.

B. Ziliotto (2013),
Zero-sum repeated games: counterexamples to the existence of the asymptotic value and the conjecture maxmin=$\lim v_n$,
hal- 00824039.